Predicting sexual and nonsexual recidivism in a consecutive sample of juveniles convicted of sexual offences

Aebi, M ; Plattner, B ; Steinhausen, H C ; Bessler, C

Abstract: Reliable and valid risk assessments are essential for responding adequately to juveniles who have sexually offended (JSO). Given the lack of specific research focusing on the JSO population, the present study aims at confirming and expanding previous findings based on clinical samples dealing with risk assessments of JSO. The predictive power of the Juvenile Sex Offender Assessment Protocol (J-SOAP-II) and the Sexual Offence Severity (SOS) Scale are evaluated retrospectively by analyzing forensic, police, and judicial files in a consecutive sample of 223 adolescents (mean age of 15.7 years, SD = 2.1 years) who had been convicted of a sexual crime in the Canton of Zurich, Switzerland, between 2000 and 2008. Based on local official recidivism data (mean follow-up period = 4.3 years; SD = 2.5 years), univariate and multivariate predictions of sexual and nonsexual recidivism are tested by use of the J-SOAP-II and the SOS Scale in logistic regression and receiver operating characteristics (ROC) analyses. Sexual recidivism (n = 7, 3.1%) is best predicted by a multivariate model, including the J-SOAP-II Adjustment Scale and the SOS Scale (ROC: Area under the curve [AUC] = .818). Nonsexual violent (n = 37, 16.6%) and general recidivism (n = 100, 44.8%) are only moderately predicted by the J-SOAP-II impulsive/antisocial behavior scale (AUC = .677 and .662, respectively). In conclusion, the J-SOAP-II adjustment scale, the J-SOAP-II impulsive/antisocial behavior scale, and the SOS Scale may be helpful for screening purposes in JSO, but additional risk assessment seems necessary. Overall, further research is necessary for a comprehensive understanding of risk factors in JSO.

DOI: https://doi.org/10.1177/1079063210384634

Predicting Sexual and Non-Sexual Recidivism in a Consecutive Sample of Juveniles Convicted of Sexual Offences

Marcel Aebi[1], Belinda Plattner[1], Hans-Christoph Steinhausen [1,2,3] and Cornelia Bessler[1]

[1] Department of Child and Adolescent Psychiatry, University of Zurich, Switzerland

[2] Aalborg Psychiatric Hospital, Aarhus University Hospital, Denmark

[3] Child and Adolescent Clinical Psychology, Institute of Psychology, University of Basel, Switzerland

Correspondence to:

Marcel Aebi

University of Zurich

Department of Child and Adolescent Psychiatry

Division of Child and Adolescent Forensic Psychiatry

Neptunstrasse 60, 8032 Zurich

Switzerland

Phone: +41 43 556 40 13

Fax: +41 43 556 40 05

E-Mail: maebi@ppkj.uzh.ch

Abstract

Reliable and valid risk assessments are essential for responding adequately to juveniles who have sexually offended (JSO). Given the lack of specific research focussing on the JSO population, the present study aimed at confirming and expanding previous findings based on clinical samples dealing with risk assessments of JSO. The predictive power of the Juvenile Sex Offender Assessment Protocol (J-SOAP-II) and the Sexual Offence Severity (SOS) scale were evaluated retrospectively by analysing forensic, police and judicial files in a consecutive sample of 223 adolescents (mean age of 15.7 years, *SD* = 2.1 years) who had been convicted of a sexual crime in the Canton of Zurich, Switzerland, between 2000 and 2008. Based on local official recidivism data (mean follow-up period 4.3 years; *SD* 2.5 years), univariate and multivariate predictions of sexual and non-sexual recidivism were tested by use of the J-SOAP-II and the SOS scale in logistic regression and receiver operating characteristics (ROC) analyses. Sexual recidivism (n = 7; 3.1%) was best predicted by a multivariate model including the J-SOAP-II Adjustment scale and the SOS scale (ROC: Area under the curve; AUC = .818). Non-sexual violent (n = 37; 16.6%) and general recidivism (n = 100; 44.8%) were only moderately predicted by the J-SOAP-II Impulsive / Antisocial Behavior scale (AUC .677; AUC .662). In conclusion, the J-SOAP-II Adjustment scale, the J-SOAP-II Impulsive / Antisocial Behavior scale and the SOS scale may be helpful for screening purposes in JSO but additional risk assessment seems necessary. Overall, further research is necessary for a comprehensive understanding of risk factors in JSO.

Key Words: Juvenile Sexual Offending, Recidivism, Risk Assessment.

.

Introduction

Both in the USA (United States Department of Justice, 2009) and the UK (United Kingdom Home Office, 1998), approximately 20% of sexual assaults have been committed by juveniles under the age of 18. Comparable data have been found for sexual assaults in Switzerland (Swiss Federal Institute for Statistics, 2009). Thus, sexual crimes committed by juveniles represent a serious problem in society. Quite often adequate responding to juveniles who have sexually offended (JSO) is a major challenge for the judicial and social systems. In the interest of public safety and reduction of the risk of re-offending, courts have to decide on incarceration, detention placement, and treatment responses to JSO. Furthermore, in some states JSO perceived to be at high risk of violent re-offending are placed on public sex offender registries (Caldwell & Dickinson, 2009). Considering the potentially large impact of these interventions, courts often rely on clinical expertise for discriminating high from low risk JSO. Hence, for forensic mental health professionals the understanding of the factors that contribute to sexual and violent re-offending is imperative.

In the recent past, the knowledge of potential risk factors for violent and sexual re-offending has grown continuously (Quinsey, Harris, Rice, & Cormier, 2006; Worling & Långström, 2006). Whereas sexual deviancy and criminal lifestyle characteristics have been identified as the major domains associated with persistent sexual offending in adults (see review of Hanson & Morton-Bourgon, 2005), factors contributing to sexual re-offending in juveniles are less clear (Caldwell, 2002; Vitacco, Caldwell, Ryba, Malesky, & Kurus, 2009). However, support has been found for some risk factors in juveniles (Worling & Långström, 2006), namely, (1) deviant sexual interest, (2) prior criminal sanctions for sexual offending, (3) sexual offending against multiple victims, (4) sexual offending against a stranger, (5) social isolation, and (6) failure to participate in specialized treatments of sexual offending. Furthermore, a number of other relevant risk factors like sexual offending against a child

victim or sexual preoccupation have been identified in clinical practice but still await empirical confirmation (Worling & Långström, 2006).

Given the insufficient prediction of sexual recidivism based on unstructured clinical judgments (Hanson & Morton-Bourgon, 2009; Quinsey et al., 2006), structural risk assessment instruments accounting for various risk factors have been developed and have been recommended for clinical use (e.g. Borum, Bartel, & Forth, 2003; Prentky & Righthand, 2003; Worling, 2004). The Juvenile Sexual Offender Assessment Protocol (J-SOAP; Prentky & Righthand, 2001) and its revision (J-SOAP-II; Prentky & Righthand, 2003) represent structural risk assessment instruments designed for juveniles adjudicated for sexual contact offences as well as non-adjudicated juveniles with a history of sexual coercive behavior between 12 and 18 years of age. The J-SOAP-II consists of 28 items that result from an extensive review of pertinent risk factors as described in the adolescent and adult sex and non-sex offender literature (Prentky & Righthand, 2001). The J-SOAP-II is conceptualized in four narrowband scales which have been organized by their content in terms of face validity. The scales are labelled (1) Sexual Drive/Preoccupation (J-SOAP-II Sexual Drive), (2) Antisocial Behavior/Impulsivity (J-SOAP-II Antisocial), (3) Intervention Scale (J-SOAP-II Intervention) and (4) Community Stability/Adjustment (J-SOAP-II Adjustment). Furthermore, two broadband scales labelled static summary scales (consisting of scale 1 and 2) and dynamic summary scales (consisting of scale 3 and 4) and a total score may be calculated. The constructive validity of the previous version of the J-SOAP-II (based on 26 items) was tested by Principal Component Analyses (PCA) and resulted in a four factor solution similar to the four narrowband scales (Righthand et al., 2005). Furthermore, adequate psychometric properties have been described (Martinez, Flores, & Rosenfeld, 2007; Parks & Bard, 2006; Prentky & Righthand, 2003). Although the J-SOAP-II and its earlier version, the J-SOAP, appear to be fairly widely used (Burton, Smith-Darden, & Frankel, 2006), relatively little is known about the predictive validity of the scales (Caldwell, 2010; Viljoen et al., 2008).

Furthermore, studies that have specifically addressed the predictive validity of the J-SOAP-II have reported rather conflicting findings. The validity of the J-SOAP-II has been partly supported in four studies. A retrospective 7 year follow-up study including a large sample of 822 JSO assigned to placement decisions by the department of social services found adequate prediction of sexual recidivism by the J-SOAP-II total score using receiver operating characteristics (ROC) analyses as the measure of accuracy (Area under the Curve; AUC = .81) (Prentky, 2006). In this same study, three narrowband scales also achieved similar AUCs, while the J-SOAP-II Antisocial scale failed to predict recidivism (AUC = .57). In a second study, the J-SOAP-II total score again predicted sexual recidivism above chance levels (AUC = .86), this time with a smaller sample of 60 urban mostly minority male JSO (Martinez et al., 2007). Furthermore, Martinez and colleagues (2007) found that the J-SOAP-II total score was also effective in predicting general recidivism (AUC = .76). Support of the validity of the dynamic scales but less for the static scales was found in additional logistic regression analyses. However, the J-SOAP-II Sexual Drive scale was not at all correlated with sexual recidivism ($r = 0.04$ $p > .05$). In a third study, Caldwell and Dickinson (2009) reported that the J-SOAP-II Antisocial scale predicted general, any violent, any sexual and violent sexual offenses (AUC's = .65, .70, .70, and .68, respectively) whereas the J-SOAP-II Sexual Drive scale did not show AUC above .5 for these offenses (AUC's = .39, .41, .47, and .50, respectively) with a sample of 172 JSO incarcerated in secured settings. Likewise, Parks and Bard (2006) found some support for the J-SOAP-II Antisocial scale for predicting sexual recidivism in 56 JSO placed in a secure correctional facility using Cox regression analyses. However, due to the multivariate prediction model including the Psychopathy Checklist: Youth Version (Hare, 2003) the unique contribution of the J-SOAP-II Antisocial scale in predicting sexual recidivism remained unclear.

However, findings from two other studies were less supportive. Based on data from a sample of 169 male JSO in a residential treatment program, Vilojen and colleagues (2008) reported

poor predictive validity of the J-SOAP-II for sexual recidivism (J-SOAP-II total score AUC = .54; narrowband scales AUC = .45 to .60), whereas the predictive validity for serious non-sexual violence was slightly better (J-SOAP-II total score AUC = .63 and narrowband scale scores from AUC = .52 to .67). Furthermore, in a study by Caldwell and colleagues (2008) the J-SOAP-II Sexual Drive scale and Antisocial scale failed to predict sexual and non-sexual recidivism in a sample of 91 JSO from a secured correctional treatment program. However, the J-SOAP-II Intervention scale did have some predictive utility for felony sexual offense charges ($R^2 = .029$, $p < .05$).

In general, research on risk assessment in JSO is complicated by the great heterogeneity in this population (Andrade, Vincent, & Saleh, 2006; Bessler, 2008; Fehrenbach, Smith, Monastersky, & Deisher, 1986; Lambie & Seymour, 2006; Ryan, Miyoshi, Metzner, Krugman, & Fryer, 1996), the rapid change due to developmental processes and social influences in this critical period of life (Prentky & Righthand, 2003; Price, 2003), and the rather low rates of sexual recidivism in JSO (Caldwell, 2010; Gerhold, Brown, & Beckett, 2007). Furthermore, research findings based on selected clinical samples could be misleading because more impaired JSO may be overrepresented so that the perceived risk for sexual re-offending may be inflated (Kjellgren, Wassberg, Carlberg, Langstrom, & Svedin, 2006). Inconsistent findings regarding the validity of the J-SOAP-II may also be due to different sampling strategies (low risk vs. high risk JSO samples), different methods in the assessment of recidivism (charges vs. social service information), and criterion contamination (lacking independence of risk and recidivism assessment). Considering the various methodological difficulties, the J-SOAP-II has not been recommended as a specialized risk assessment instrument for JSO (Vitacco et al., 2009). Consequently, the authors of the J-SOAP-II emphasize the current limitations of the J-SOAP-II and recommend the use of additional information sources in order to arrive at the final decision on risk status (Prentky & Righthand, 2003). Because the J-SOAP-II

requires additional validation, further research based on clinical and criminal population samples is highly needed (Prentky & Righthand, 2003).

To overcome some of the constraints mentioned above, the present study aimed at testing the J-SOAP-II in the prediction of sexual, non-sexual violent, and general recidivism in a consecutive sample of convicted JSO that is more representative than selected clinical samples. Furthermore, due to the insufficient validity of the J-SOAP-II Sexual Drive scale (Caldwell et al., 2008; Martinez et al., 2007; Viljoen et al., 2008), we included two additional sexual offence measures in the present study. First, a modified version of the sexual severity scale (Aylwin et al., 2000) describing the sexual intensity of the offence characteristics was also analyzed in the prediction of sexual recidivism. The degree of contact by the offender has been considered in risk assessment instruments for sex offender registration (Juvenile Risk Assessment Scale, JRS; Codey & Harvey, 2007). Second, as an additional measure of sexual preoccupation, the number of sexually abusive acts against the assault victim(s) was evaluated using court file information.

Methods

Sample description

The present study was based on a retrospective analysis of the forensic, police and judicial files and was designed in collaboration with the justice department of the Canton of Zurich, Switzerland. In addition to police reports and court protocols, every case file contained at least a short summary by a social worker on social and family status, school experiences, and a description of potential behavioral and familial difficulties of the young person. In addition, 90 of the 223 JSO included in the final sample had been assessed by a forensic mental health professional. The study was approved by the local medical ethics committee. The original sample included all children and adolescents aged between 10 and 18 years

who had been convicted of a sexual assault against children (victims under 16 years and at least 3 years younger than the offender), coercive sexual behaviour, rape, exhibitionism, sexual harassment, porn distribution or illegal pornography in the Canton of Zurich between 2000 and 2008. A total number of 419 JSO had been reported. However, not all of these juveniles had been finally convicted; 40 (9.5%) of the youngsters were not found to be guilty by a superior court. In the present study, the juveniles exclusively convicted of porn distribution or illegal pornography (n = 119; 28.4%) were excluded from further analyses. Given that the J-SOAP-II was designed for use with male youth and given that risk factors may differ between males and females, the 9 (2.1%) female JSO were excluded. From the remaining 251 subjects, 28 (11.2%) files were not available, mostly due to official elimination of archive files (*n*= 22, 8.8%). Attrition analyses showed that the 28 cases with missing file information did not differ from the 223 youngsters in the final sample in terms of recidivism rate (sexual recidivism [10.7 % vs. 3.1%, χ^2 = 3.73, *df* = 1, *p* > *.05*], non-sexual violent recidivism [10.7% vs. 16.6%, χ^2 = .64, *df* = 1, *p* > *.05*], general recidivism [64.3 % vs. 44.8%, χ^2 = 3.78, *df* = 1, *p* > *.05*]). However, the 28 cases with missing file information were significantly younger than the remaining sample of 223 participants (14.50 vs. 15.64, *t* = -2.37, *df* = 32.68, *p* < 0.05). The final sample had a mean age of 15,7 years (*SD* = 2.1 years).

Assessment

The present study is based on retrospective evaluations of sexual offenses as documented in forensic, police and judicial files. In some files the information on the JSO intervention and treatment was limited. Data were extracted by an experienced forensic expert and two master students with bachelor degrees in psychology. In addition, outcome data were collected from the local crime registry. The coders did not know the participants' recidivism status.

Juvenile Sexual Offender Assessment Protocol revised (J-SOAP-II)

The J-SOAP-II is a 28-item checklist of risk factors that are assembled in four independent narrowband scales (Sexual Drive/Preoccupation, Antisocial Behavior/Impulsivity, Intervention, Community Stability/Adjustment), two broadband scales (static summary and dynamic summary scales), and a total score. The scales and the total score were designed to assess risks for sexual violence as well as general delinquency. The instrument is intended for use specifically with adolescent boys aged between 12 and 18 years who have a history of sexually coercive behaviour (Prentky & Righthand, 2003). Items are scored on a 3-point scale in which a score of 0 is associated with the absence of the risk factor and a score of 2 indicates clear evidence that the factor is present. A score of 1 implies the presence of some information that suggests the presence of the item but the information is insufficient or unclear to justify a score of 2. For some items a score of 1 refers to the degree, intensity or count of the item. When the information is limited, unclear, or incomplete, raters are instructed to score items in the direction of lower risk (Prentky & Righthand, 2003). The present study was based on the German version of the J-SOAP-II (Schmelzle, 2004). Three trained assessors coded the J-SOAP-II items entirely based on the data in the forensic files.

Sexual Offence Severity Scale

The development of the Sexual Offence Severity (SOS) scale was based on previous research on the severity of sexual abuse (e.g. Firestone, Bradford, Greenberg, & Nunes, 2000; Koss & Dinero, 1988) and was originally designed for comparing adult and adolescent sexual offenders by Aylwin and colleagues (2000). The scale provides a framework for describing offence severity in isolation of other factors. Thus, methods of sexual assault have been ranked by intensity and aggression criteria irrespective of the effect upon the victim (Aylwin et al., 2000). The most severe sexual assault by an offender has to be coded on a rating scale describing specific offence characteristics on six levels: (1) Severity level one including clothed fondling, voyeurism and obscene phone calls, (2) severity level two

including off-clothed fondling, digital penetration, masturbation, exhibitionism and frotteurism, (3) severity level three including oral sex and simulated intercourse, (4) severity level four including attempted or performed vaginal intercourse, (5) severity level five including attempted or performed anal intercourse and victim gang rape, and (6) severity level six including offence levels of particular brutality (e. g. forced confinement, humiliation and weapon use).

Based on the file report the most severe sexual assault of an offender was coded in the present study. The highest severity level was excluded from the present study because humiliation and extreme force could not be measured reliably. Furthermore, severity level one was split up in less severe non-contact (e.g. obscene phone calls) and more severe contact offences (fondling). Thus, the final SOS scale consisted of six levels again.

Number of sexual assaults against victim(s)

The number of sexual offences committed against the victim(s) by the offender during the index offense was computed, based on file information using the final judgement of the court files. Specifically, more than one assault was counted when the offender had left the victim for at least one hour and then reoffended against the victim. Sexual assaults from earlier convictions or adjudications were not included because this information is captured by the J-SOAP-II Sexual Drive scale.

Descriptive file information

To describe various recidivism groups, victim characteristics (male victim, child victim, multiple victims, family victim, unknown victim) and offender characteristics (group offender, repeated offender and use of alcohol or drugs) were identified from the file data.

Outcome measures

Recidivism information on adolescent and adult charges came from the local crime registry of the Canton of Zurich and was collected after the final administration of the J-SOAP-II. The mean time for recidivism observation was 4.30 years (*SD* = 2.49) with a range of 9 months to 9.74 years. The computerized data base contains all past and current transactions from all prosecution institutions and prisons in the Canton of Zurich including the information on the date of the charges, the type of the offence, date of convictions or penalty orders, and the beginning and end of detentions or incarcerations. The data base is limited to data obtained in the Canton of Zurich. As an additional limitation, the data base does not contain sentence or court information. Therefore, we use charges only and not convictions for recidivism measure in the present analyses.

Statistical analyses

First, descriptive analyses were performed on victim and offence characteristics. Second, interrater reliability was calculated by using Intraclass correlation coefficients (ICC) in terms of the two-way random based approach on a single measure [ICC(2.1)] (Shrout & Fleiss, 1979) for three independent raters with similar training experiences using a random sub-sample of 60 cases. Third, Cronbach's alpha as a measure of internal consistency was calculated. Finally, to study the accuracy in the prediction of (1) sexual, (2) non-sexual violent, and (3) general recidivism, ROC analyses were performed separately for each J-SOAP-II narrowband scale, the two broadband scales, and the total score. In addition, the SOS scale, the number of sexual offences, and a multivariate prediction model were included in the ROC analyses. In ROC analyses, all possible combinations of sensitivity and specificity that can be achieved by changing the test's cut-off score can be summarized by use of a single parameter called the area under the ROC curve (AUC) (for additional information on ROC analyses see Streiner & Cairney, 2007). According to Ferdinand (2008),

the AUC as a measure of excellence for predicting recidivism should be interpreted as follows: poor (50-.70); fair (.71-.80); good (.81-.90), and excellent (.91-1.00).

A two step procedure was applied in the development of multivariate recidivism prediction models. First, univariate logistic regression (LR) analyses for each J-SOAP-II narrowband scale, the SOS scale, and the number of sexual offences were performed separately. In general, the LR is an exploratory statistical method and therefore will more readily result in significant findings than ROC analyses. Secondly, those scales which were significant predictors in univariate LR were used as predictors in subsequent backward LR analyses (removal criterion = .20) in order to select a combination of predictors of either sexual, non-sexual violent, or general recidivism. If a significant multivariate model was identified on the basis of the final LR model, the probabilities for each subject were calculated and subjected to ROC analyses in order to compare the accuracy of the multivariate model with the original scales and measures.

Results

Descriptive findings

The present sample of JSO had been convicted of the following assaults: Sexual abuse of a child (*n* = 95, 42.6%), sexual coercion or rape (*n* = 125, 56.1%), sexual harassment (*n* = 46, 20.6%) or exhibitionism (*n* = 8, 3.6%). The percentages did not sum up to 100, because various JSO may had been convicted of multiple sexual crimes. A total of 202 (90.6%) had received a juridical order in form of a penalty (7.6% forfeit, 40.0% obligated to work, 6.3% incarceration, 4.5% other) and/or a sanction (22.4% personal supervision, 34.1% outpatient therapy, 17.0% placement in foster care or youth institutions, 0.4% psychiatric hospitalization) whereas in 3 (1.3%) of the subjects no information was available.

Descriptive statistics for the three groups of recidivism

In the follow-up period 7 JSO did re-offend sexually, 37 did show non-sexual violent re-offending, and 100 of the 223 JSO did show general recidivism (including misdemeanours). Table I presents victim and offender characteristics of re-offenders and non-re-offenders by recidivism group. JSO with a child victim under the age of 10 show significantly fewer non-sexual violent ($\chi^2 = 5.38$, $p < .05$) and general ($\chi^2 = 5.61$, $p < .05$) recidivism compared to JSO with no victims under 10 years of age. In addition JSO with a family victim (siblings other related persons) show less general recidivism ($\chi^2 = 4.28$, $p < .05$) than JSO with no family victim. The three recidivism groups did not differ significantly with respect to the remaining victim characteristics (male victim, multiple victims, and unknown victims) or offender characteristics (group offence, repeated offences and the use of alcohol or drugs).

Reliability analyses of the predictors

Table II shows the means, standard deviations, ranges, ICC, and Cronbach`s alpha coefficients of the J-SOAP-II total score, the J-SOAP-II static and dynamic summary scales,

and the four J-SOAP-II narrowband scales. The ICC for the two narrowband (J-SOAP-II Antisocial and Adjustment) and one broadband scale (J-SOAP-II static) were below .60 whereas the ICC for the remaining JSOAP-II scales showed ICC between .67-.71. The variable representing the number of sexual assaults against index victim(s) showed high interrater reliability (ICC = .95 for three raters). Lastly, Kendalls W as a measure for the interrater reliability of the ordinal SOS scale was .95. Cronbach's alpha ranged from .76 to .87 for all J-SOAP-II scales except the J-SOAP-II Sexual Drive scale (α = .56).

Univariate prediction of recidivism

Table III shows the odds ratios from univariate LR and AUC from ROC analyses for the J-SOAP-II scales and total score, the number of sexual assaults against index victim(s) and the SOS scale for the prediction of sexual, non-sexual violent, and general recidivism. Using univariate LR the J-SOAP-II total score, J-SOAP static, J-SOAP-II Antisocial scale, J-SOAP-II Adjustment scale, and the SOS scale were significantly associated with sexual recidivism (OR ranging between 1.23 and 1.86) Furthermore, ROC analyses revealed that sexual recidivism was significantly predicted by J-SOAP-II Antisocial scale (AUC .739), J-SOAP-II Adjustment scale (AUC = .743) and the SOS scale (AUC = .751) but not by the remaining J-SOAP-II scales or the number of sexual assaults against the index victim(s).

Furthermore, all J-SOAP-II scales except the J-SOAP-II Sexual Drive scale were significant predictors of non-sexual violence using univariate LR and ROC analyses (Table III). The OR of these predictors ranged between 1.05 and 1.19 whereas AUC ranged between .603 and .677. Finally, all J-SOAP-II scales except the J-SOAP-II Sexual Drive scale significantly predicted general recidivism in the LR analysis. (OR ranged between 1.05 and 1.19) and all J-SOAP-II scales except the Sexual Drive and Intervention scales predicted general recidivism in the ROC analyses (Table III).

Multivariate prediction of recidivism

The J-SOAP-II Antisocial, J-SOAP-II Adjustment and the SOS scale were further analyzed in multivariate prediction models of sexual recidivism. Furthermore, the J-SOAP-II Antisocial, J-SOAP-II Intervention and J-SOAP-II Adjustment scales were included in multivariate prediction of general recidivism. Because the J-SOAP-II Antisocial scale was the only significant predictor of non-sexual violent recidivism, no further multivariate analysis was feasible. Table IV shows the final results from the multivariate LR predicting sexual and general recidivism. A significant two factor prediction model ($\chi^2 = 10.51$, $df = 2$, $p < 0.01$) was found for sexual recidivism including the J-SOAP-II Adjustment and SOS scales whereas for the multivariate prediction of general recidivism only the J-SOAP-II Antisocial scale was found to be significant. On the basis of the final LR model, the probabilities for each subject were calculated and subjected to ROC analyses.

Discussion

General findings

In agreement with previous findings (Caldwell, 2010; Gerhold et al., 2007; Worling & Långström, 2006), sexual recidivism was low in the present study (3.1%) whereas non-sexual violent (16.6%) and general recidivism was substantially higher (44.8%). Hence, the present study confirms previous findings that JSO are often versatile offenders who are at risk of future multiple delinquency (Caldwell, 2010). Interestingly, the analyses of the offence data suggest that sexual assault of a child was negatively associated with further violent and general recidivism. Accordingly, general delinquency may be less pronounced in a particular subsample of JSO.

Reliability of the J-SOAP-II

According to the interpretation of reliability scores by Landis and Koch (1977), the interrater reliabilities of the J-SOAP-II scales ranged from moderate (.41-.60) to substantial (.61-.80). Three scales (the narrowband J-SOAP-II Antisocial and Adjustment scales and the broadband J-SOAP-II Static scale) showed ICC scores below .60. Internal consistency was acceptable (Cronbach's alpha > .75) for all J-SOAP-II scales with one exception, namely, the Sexual Drive scale with a Cronbach's alpha of just .56. Overall, the reliability of the J-SOAP-II coded from file information was limited tough still sufficient for further analyses.

Prediction of sexual recidivism

Perhaps the most relevant task of the J-SOAP-II is to predict sexual recidivism. However, according to ROC results only the J-SOAP-II Antisocial (AUC = .739) and the J-SOAP-II Adjustment (AUC = .743) scales but not the J-SOAP-II Sexual Drive, total score or summary scales were significant predictors of sexual recidivism. In the logistic regression analyses additional predictors were significant (J-SOAP-II total score and the static summary score). However, the OR resulting from these LR models were rather low and disappointing. Overall, an acceptable predictive validity was found only for the J-SOAP-II Antisocial and the J-SOAP-II Adjustment scale. Thus, in agreement with previous findings (Caldwell & Dickinson, 2009; Caldwell et al., 2008; Martinez et al., 2007; Viljoen et al., 2008), the J-SOAP-II Sexual Drive Scale was not a significant predictor of sexual re-offence. In many other respects, present results differed from the existing literature. For example, previous studies have identified the J-SOAP-II total score and the J-SOAP-II Intervention scale as valid predictors of sexual recidivism (Martinez et al., 2007; Prentky, 2006) but that was not the case in the present study. Likewise, whereas the present study identified the J-SOAP-II Antisocial and Adjustment scales as predictive of sexual recidivism, other studies have failed to do so (Caldwell et al., 2008; Viljoen et al., 2008). The present study also found that the SOS scale

was predictive of sexual recidivism, and improved multivariate outcomes when included with the J-SOAP-II Adjustment scale.

Different sample characteristics may be responsible for some of the inconsistent findings regarding the prediction of sexual recidivism in JSO. Previous studies have focused on JSO assigned for treatment purposes (Caldwell et al., 2008; Martinez et al., 2007; Parks & Bard, 2006; Prentky, 2006; Viljoen et al., 2008; Waite et al., 2005) while the present analyses were based on a consecutive criminal cohort of JSO. As a consequence, the present sample may be more heterogeneous and not comparable to selected clinical samples regarding offence characteristics and personality of the JSO.

Interestingly, the predictive value of the SOS scale was in contrast to a previous study which found that any form of penetration was associated with a decrease of sexual re-offending (Långström, 2002). However, this study was restricted to 15 to 20 years old adolescents who had been mostly convicted of rape (73%). Thus, this finding may be not comparable to the present study which involved younger JSO and a broader spectrum of sexual assaults. In addition, the SOS scale may not function as a predictor of sexual recidivism in clinically referred JSO because this subgroup might be already preselected by offence severity criteria. Moreover, some sort of non-contact offences such as exhibitionistic acts, which often appear in combination with deviant sexual interests, have shown high rates of sexual recidivism in JSO (Långström & Grann, 2000). Combining exhibitionism and other sexual offences may conceal important information on sexual offence severity as a risk factor (Worling & Långström, 2006). Because only eight cases with exhibitionism were included in the present sample, this finding may not have biased the present results very much.

Prediction of non-sexual violence and general recidivism

The J-SOAP-II total score and some narrowband scales significantly predicted non-sexual violence and general recidivism in the present sample. However, for all the J-SOAP-II scales, the AUC from ROC and the OR from LR were only moderate so that their predictive power seems to be limited. In agreement with previous findings (Caldwell et al., 2008; Viljoen et al., 2008; Waite et al., 2005) and theoretical considerations (Prentky & Righthand, 2001, 2003), the J-SOAP-II Antisocial scale received the highest values for non-sexual violent (AUC = .677, OR 1.20) and general recidivism (AUC = .662, OR 1.19).

Limitations

Because the present study is expanding the use of the J-SOAP-II to a broader spectrum of JSO, the present findings may be neither comparable to community treatment sample nor to residential treatment samples, which may under- or over-represent the risk of JSO. Furthermore, due to the specific sample characteristics the results may be limited to male Caucasian JSO living in urbanized areas. In addition, further limitations are noteworthy. First, all information has been extracted retrospectively based on the available file information. Low frequencies of particular items may have had an impact on reliability of some of the scales used in the present study. A prospective design and a direct forensic assessment of the JSO may improve data quality in future studies. In particular, the information for the J-SOAP-II Intervention scale was incomplete, which may have reduced the predictive validity of this scale. However, most of the information has been coded reliablely according to the guidelines of the J-SOAP-II manual (Prentky & Righthand, 2003). Secondly, the evaluation of recidivism was based on local official data only and did not include national official data or self reports so that there may be some under-reporting. Thirdly, like most studies dealing with sexual recidivism in JSO (Caldwell, 2010; Gerhold et al., 2007; Prentky & Righthand, 2003; Worling & Långström, 2006), we found a rather low base rate for sexual recidivism (3.0%). The small sub-sample of JSO with sexual re-offending decreased the statistical

power and hindered further attempts to validate and to improve risk assessment in JSO. In other words, because of the small absolute number of recidivists in the present study, a few cases of undetected sexual recidivism may have had a strong impact on the present findings.

Conclusions

Despite these limitations, the present study may stimulate research on risk assessment in JSO. The study of representative criminal samples may improve the predictive power of risk assessment instruments for JSO. Evaluators are often asked to conduct risk assessments prior to treatment or placement decisions. The samples analysed in the present study include the full range of youth who might face such an evaluation. Therefore, the present results may improve important clinical decision making processes. In conclusion, the J-SOAP-II Antisocial and the J-SOAP-II Adjustment scale may be helpful for screening purposes of convicted JSO with a widespread range of sexual misbehaviors. However, given the moderate reliability of these scales, some additional risk assessment may be necessary. The sexual offence severity should be considered for risk assessment and intervention planning. JSO scoring high in one of these scales should be further assessed by forensic mental health professionals using additional risk instruments for violent and general offending. Up to now, it remains unclear which sexual behavioural patterns are associated with sexual recidivism. However, based on the present and previous findings (Caldwell & Dickinson, 2009; Caldwell et al., 2008; Martinez et al., 2007; Viljoen et al., 2008) the J-SOAP-II Sexual Drive scale did not adequately address sexual recidivism and should, therefore, this score should be excluded from risk assessments by clinicians. Finally, research on the J-SOAP-II may be further specified by abstaining from a total score which is simply based on a sum of isolated risk factors. By creating a more complex algorithm accounting for different JSO subtypes (Butler & Seto, 2002; Hunter, Figueredo, Malamuth, & Becker, 2003), the prognosis of JSO recidivism may be improved. In addition, different scale scores will provide most likely

more useful information not only for determining risk but also for disposition and treatment planning.

Acknowledgement

The authors are grateful to Karl Brühwiler and Gunnar Vogt for assisting with data extraction from the files.

References

Andrade, J. T., Vincent, G. M., & Saleh, F. M. (2006). Juvenile sex offenders: a complex population. *Journal of Forensic Science, 51*(1), 163-167.

Aylwin, A. S., Clelland, S. R., Kirkby, L., Reddon, J. R., Studer, L. H., & Johnston, J. (2000). Sexual offense severity and victim gender preference: A comparison of adolescent and adult sex offenders. *International Journal of Law and Psychiatry, 23*(2), 113-124.

Bessler, C. (2008). Jugendliche Sexualstraftäter, Persönlichkeitsmerkmale, Beurteilungsverfahren und Behandlungsansätze (Juvenile sexual offenders, personal characteristics, assassment and treatment). In H.-C. Steinhausen & C. Bessler (Eds.), *Jugenddelinquenz: Entwicklungspsychiatrische und forensische Grundlagen und Praxis (Juvenile Delinquency: Developmental Psychiatric and Forensic Basics and Practice)* (pp. 176-199). Stuttgart: Kohlhammer.

Borum, R., Bartel, P., & Forth, A. (2003). *Manual for the Structured Assessment of Violence Risk in Youth, Version 1.1*: University of South Florida.

Burton, D., Smith-Darden, J., & Frankel, S. (2006). Research on adolescent sexual abuser treatment programs. In H. E. Barbaree & W. L. Marshall (Eds.), *The Juvenile Sex Offender* (2nd ed., pp. 291-312). New York: Guilford.

Butler, S. M., & Seto, M. C. (2002). Distinguishing two types of adolescent sex offenders. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(1), 83-90.

Caldwell, M. F. (2002). What we do not know about juvenile sexual reoffense risk. *Child Maltreatment, 7*(4), 291-302.

Caldwell, M. F. (2010). Study characteristics and recidivism base rates in juvenile sex offender recidivism. *International Journal of Offender Therapy and Comparative Criminology, 54*(2), 197-212.

Caldwell, M. F., & Dickinson, C. (2009). Sex offender registration and recidivism risk in juvenile sexual offenders. *Behavioral Science & the Law, 27*(6), 941-956.

Caldwell, M. F., Ziemke, M., & Vitacco, M. (2008). An examination of the Sex Offender Registration and Notification Acts as applied to juveniles: Evaluation the ability to predict sexual recidivism. *Psychology, Public Policy, and Law, 14*, 89-114.

Codey, R., & Harvey, P. (Eds.). (2007). *Attorney general guidelines for law enforcement for the implementation of sex offender registration and community notification laws.*

Fehrenbach, P. A., Smith, W., Monastersky, C., & Deisher, R. W. (1986). Adolescent sexual offenders: Offender and offense characteristics. *American Journal of Orthopsychiatry, 56*(2), 225-233.

Ferdinand, R. F. (2008). Validity of the CBCL/YSR DSM-IV scales anxiety problems and affective problems. *Journal of Anxiety Disorders, 22*(1), 126-134.

Firestone, P., Bradford, J. M., Greenberg, D. M., & Nunes, K. L. (2000). Differentiation of homicidal child molesters, nonhomicidal child molesters, and nonoffenders by phallometry. *American Journal of Psychiatry, 157*(11), 1847-1850.

Gerhold, C. K., Brown, D. B., & Beckett, R. (2007). Predicting recidivism in adolescent sexual offenders. *Aggression and Violent Behavior, 12*, 427-438.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*, 839-843.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: a meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*(6), 1154-1163.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: a meta-analysis of 118 prediction studies. *Psychological Assessment, 21*(1), 1-21.

Hare, R. D. (Ed.). (2003). *The Hare psychopathy checklist - revised*. Toronto: Multi-Health Systems.

Hunter, J. A., Figueredo, A. J., Malamuth, N. M., & Becker, J. V. (2003). Juvenile sex offenders: Toward the development of a typology. *Sexual Abuse, 15*(1), 27-48.

Kjellgren, C., Wassberg, A., Carlberg, M., Langstrom, N., & Svedin, C. G. (2006). Adolescent sexual offenders: A total survey of referrals to Social Services in Sweden and subgroup characteristics. *Sexual Abuse, 18*(4), 357-372.

Koss, M. P., & Dinero, T. E. (1988). Predictors of sexual aggression among a national sample of male college students. *Annals of the New York Academy of Sciences, 528*, 133-147.

Lambie, I., & Seymour, F. (2006). One size does not fit all: Future directions for the treatment of sexually abusive youth in New Zealand. *Journal of Sexual Aggression, 12*(2), 175-187.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Långström, N. (2002). Long-term follow-up of criminal recidivism in young sex offenders: Temporal patterns and risk factors. *Psychology, Crime & Law, 8*, 41-58.

Långström, N., & Grann, M. (2000). Risk for criminal recidivism among young sex offenders. *Journal of Interpersonal Violence, 15*(8), 855-871.

Martinez, R., Flores, J., & Rosenfeld, B. (2007). Validity of the juvenile sex offender assessment protocol-II (J-SOAP II) in a sample of urban minority youth. *Criminal Justice and Behavior, 34*(10), 1284-1295.

Parks, G. A., & Bard, D. E. (2006). Risk factors for adolescent sex offender recidivism: evaluation of predictive factors and comparison of three groups based upon victim type. *Sexual Abuse, 18*(4), 319-342.

Prentky, R. (2006). Risk management of sexually abusive youth: A follow up study, Available from http://www.ncjrs.gov/pdffiles1/nij/grants/214261.pdf

Prentky, R., & Righthand, S. (2001). *Juvenile Sexual Offender Assessment Protocol (J-SOAP)*. Bridgewater, MA: Justice Ressource Institute.

Prentky, R., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol II (JSOAP-II) Manual.* Washington DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Price, D. (2003). A developmental perspective of treatment for sexually abusive youth. *Sexual Addiction & Compulsivity, 10*, 225-245.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders appraising and managing risk* (2nd ed.). Washington, D.C.: American Psychological Association.

Righthand, S., Prentky, R., Knight, R., Carpenter, E., Hecker, J. E., & Nangle, D. (2005). Factor structure and validation of the juvenile sex offender assessment protocol (J-SOAP). *Sexual Abuse, 17*(1), 13-30.

Ryan, G., Miyoshi, T. J., Metzner, J. L., Krugman, R. D., & Fryer, G. E. (1996). Trends in a national sample of sexually abusive youths. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(1), 17-25.

Schmelzle, M. (2004). German version of the Juvenile Sex Offender Assessment Protocol II (JSOAP-II), Available from http://www.forio.ch/PDF/gutachten/JSOAP_ÜBERSETZUNG.pdf

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Streiner, D. L., & Cairney, J. (2007). What's under the ROC? An introduction to receiver operating characteristics curves. *Canadian Journal of Psychiatry, 52*(2), 121-128.

Swiss Federal Institute for Statistics. (2009). Criminal law. from http://www.bfs.admin.ch/bfs/portal/de/index/themen/19.html

United Kingdom Home Office. (1998). *Criminal Statistics for England and Wales 1997*. London: Home Office.

United States Department of Justice (2009). Uniform Crime Report, Available from http://www.fbi.gov/ucr/ucr.htm#cius

Viljoen, J. L., Scalora, M. J., Cuadra, L., Bader, S., Chavez, V., Ullman, D., et al. (2008). Assessing risk for violence in adolscents who have sexually offended. A comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY. *Criminal Justice and Behavior, 35*(1), 5-23.

Vitacco, M. J., Caldwell, M., Ryba, N. L., Malesky, A., & Kurus, S. J. (2009). Assessing risk in adolescent sexual offenders: Recommendations for clinical practice. *Behavioral Sciences & the Law 27*(6), 929-940.

Waite, D., Keller, A., McGarvey, E. L., Wieckowski, E., Pinkerton, R., & Brown, G. L. (2005). Juvenile sex offender re-arrest rates for sexual, violent nonsexual and property crimes: a 10-year follow-up. *Sexual Abuse, 17*(3), 313-331.

Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR): preliminary psychometric data. *Sexual Abuse, 16*(3), 235-254.

Worling, J. R., & Långström, N. (2006). Risk of sexual recidivism in adolescents who offend sexually: Correlates and assessment. In H. E. Barbaree & W. L. Marshall (Eds.), *The juvenile sexual offender* (2nd ed., pp. 219-247). New York: Guilford Press.